

What is Claimed:

1. A method for providing pre-processed data for the training of mining models from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, comprising:
 - determining at least one mining structure variable from among said set of at least one variable;
 - for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data;
 - performing mining model initial processing on said retrieved values; and
 - storing the results of said mining model initial processing.
2. The method of claim 1, where said step of determining at least one mining structure variable from among said set of at least one variable comprises:
 - accepting creation operation data comprising data comprising the identity of said mining structure variables.
3. The method of claim 2, where said at least one mining structure variable comprises a continuous variable, where said creation operation data comprises an indication regarding discretization of said continuous variable, and where said step of performing mining model initial processing on said retrieved values comprises discretizing said continuous variable according to said indication.
4. The method of claim 3, where said indication comprises an indication of a number of buckets into which said continuous variable should be discretized.
5. The method of claim 3, where said indication comprises an indication of sub-ranges into which said continuous variable should be discretized.
6. The method of claim 1, where said stored results are associated with at least one mining model, and where each of said at least one mining model is trained using said stored results.
7. A computer readable medium comprising computer executable modules having computer executable instructions, said modules providing pre-processed data for the training of mining models from data set training data comprising at least one set of case data, each of said sets of

case data comprising a stored value for at least one variable from among a set of at least one variable, said computer executable modules comprising:

- a mining structure variable determination module for determining at least one mining structure variable from among said set of at least one variable;

- a data set training data retrieval module for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data;

- an initial processing module for performing mining model initial processing on said retrieved values; and

- a storage module for storing the results of said mining model initial processing.

8. The computer readable medium of claim 7, where said mining structure variable determination module accepts creation operation data comprising data comprising the identity of said mining structure variables.

9. The computer readable medium of claim 8, where said at least one mining structure variable comprises a continuous variable, where said creation operation data comprises an indication regarding discretization of said continuous variable, and where said initial processing module discretizes said continuous variable according to said indication.

10. The computer readable medium of claim 9, where said indication comprises an indication of a number of buckets into which said continuous variable should be discretized.

11. The computer readable medium of claim 9, where said indication comprises an indication of sub-ranges into which said continuous variable should be discretized.

12. The computer readable medium of claim 9, where said stored results are associated with at least one mining model, and where each of said at least one mining model is trained using said stored results.

13. An application programming interface for use in connection with providing pre-processed data for the training of mining models from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, wherein said application programming interface receives as input creation operation data comprising data comprising the identity of mining structure variables from among said set of at least one variable; for each case, retrieves a stored value for

each of said at least one mining structure variables from said data set training data; performs mining model initial processing on said retrieved values; and stores the results of said mining model initial processing.

14. The application programming interface of claim 13, where said at least one mining structure variable comprises a continuous variable, where said creation operation data comprises an indication regarding discretization of said continuous variable, and where said application programming interface discretizes said continuous variable according to said indication.

15. The application programming interface of claim 14, where said indication comprises an indication of a number of buckets into which said continuous variable should be discretized.

16. The application programming interface of claim 14, where said indication comprises an indication of sub-ranges into which said continuous variable should be discretized.

17. The application programming interface of claim 13, wherein said query is sent and said stored results are retrieved via at least one network.

18. The application programming interface of claim 13, where said stored results are associated with at least one mining model, and where each of said at least one mining model is trained using said stored results.

19. A system for providing pre-processed data for the training of mining models from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, said system comprising:

an application programming interface, said application programming interface (a) receiving as input creation operation data comprising data comprising the identity of mining structure variables from among said set of at least one variable; (b) for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data; (c) performs mining model initial processing on said retrieved values; and (d) stores the results of said mining model initial processing; and

a database for storing said data set, operably connected with said application

programming interface, and for returning said stored values to said application programming interface.

20. A system for providing pre-processed data for the training of mining models from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, said system comprising:

determination means for determining at least one mining structure variable from among said set of at least one variable;

retrieval means for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data;

initial processing means for performing mining model initial processing on said retrieved values; and

storage means for storing the results of said mining model initial processing.

21. The system of claim 20, where said determination means comprises:

data acceptance means for accepting creation operation data comprising data comprising the identity of said mining structure variables.

22. The system of claim 21, where said at least one mining structure variable comprises a continuous variable, where said creation operation data comprises an indication regarding discretization of said continuous variable, and where initial processing means comprises discretization means for discretizing said continuous variable according to said indication.

23. The system of claim 22, where said indication comprises an indication of a number of buckets into which said continuous variable should be discretized.

24. The system of claim 22, where said indication comprises an indication of sub-ranges into which said continuous variable should be discretized.

25. The application programming interface of claim 22, where said stored results are associated with at least one mining model, and where each of said at least one mining model is trained using said stored results.

26. A method for the training of a mining model from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, comprising:

determining at least one mining structure variable from among said set of at least one variable;

for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data;

performing mining model initial processing on said retrieved values;

storing the results of said mining model initial processing in a mining structure; and

training said mining model using said stored results.

27. The method of claim 26, further comprising:

storing connection data indicating that said mining model has been trained on data from said mining structure.

28. The method of claim 26, further comprising:

accepting a drill through query for specified data from said mining structure and providing said specified data.

29. The method of claim 26, where additional mining models are associated with said mining structure, and where said method further comprises:

training each of said additional mining models using said stored results.

30. The method of claim 26, where said mining structure is treated as a first class object in a database.

31. A computer readable medium comprising computer executable modules having computer executable instructions, said modules training a mining model from data set training data comprising at least one set of case data, each of said sets of case data comprising a stored value for at least one variable from among a set of at least one variable, said modules comprising:

a mining structure variable determination module for determining at least one mining structure variable from among said set of at least one variable;

a data set training data retrieval module for each case, retrieving a stored value for each of said at least one mining structure variables from said data set training data;

an initial processing module for performing mining model initial processing on said

retrieved values;

a storage module for storing the results of said mining model initial processing; and
a training module for training a mining model using said stored results.

32. The computer readable medium of claim 31, said modules further comprising:
connection data storage module storing connection data indicating that said mining model has been trained on data from said mining structure.
33. The computer readable medium of claim 31, said modules further comprising:
drill through module for accepting a drill through query for specified data from said mining structure and providing said specified data.
34. The computer readable medium of claim 31, where additional mining models are associated with said mining structure, and where said training module further trains each of said additional mining models using said stored results.
35. The computer readable medium of claim 31, where said mining structure is treated as a first class object in a database.